

ALMOST-PRIME k -TUPLES

JAMES MAYNARD

ABSTRACT. Let $k \geq 2$ and $\Pi(n) = \prod_{i=1}^k (a_i n + b_i)$ for some integers a_i, b_i ($1 \leq i \leq k$). Suppose that $\Pi(n)$ has no fixed prime divisors. Weighted sieves have shown for infinitely many integers n that $\Omega(\Pi(n)) \leq r_k$ holds for some integer r_k which is asymptotic to $k \log k$. We use a new kind of weighted sieve to improve the possible values of r_k when $k \geq 4$.

1. INTRODUCTION

We consider a set of integer linear functions

$$(1.1) \quad L_i(x) = a_i x + b_i, \quad i \in \{1, \dots, k\}.$$

We say such a set of functions is *admissible* if their product has no fixed prime divisor. That is, for every prime p there is an integer n_p such that none of $L_i(n_p)$ are a multiple of p . We are interested in the following conjecture.

Conjecture (Prime k -tuples Conjecture). *Given an admissible set of integer linear functions $L_i(x)$ ($i \in \{1, \dots, k\}$), there are infinitely many integers n for which all the $L_i(n)$ are prime.*

With the current technology it appears impossible to prove any case of the prime k -tuples conjecture for $k \geq 2$.

Although we cannot prove that the functions are simultaneously prime infinitely often, we are able to show that they are *almost prime* infinitely often, in the sense that their product has only a few prime factors. This was most notably achieved by Chen [1] who showed that there are infinitely many primes p for which $p + 2$ has at most 2 prime factors. His method naturally generalises to show that for a pair of admissible functions the product $L_1(n)L_2(n)$ has at most 3 prime factors infinitely often.

Similarly sieve methods can prove analogous results for any k . We can show that the product of k admissible functions $\Pi(n) := L_1(n) \dots L_k(n)$ has at most r_k prime factors infinitely often, for some explicitly given value of r_k . We see that the prime k -tuples conjecture is equivalent to showing we can have $r_k = k$ for all k . The current best values of r_k grow asymptotically like $k \log k$ and explicitly for small k we can take $r_2 = 3$ (Chen, [1]), $r_3 = 8$ (Porter, [9]), $r_4 = 12$, $r_5 = 16$, $r_6 = 20$ (Diamond and Halberstam [2]), $r_7 = 24$, $r_8 = 28$, $r_9 = 33$, $r_{10} = 38$ (Ho and Tsang, [6]). Heath-Brown [5] showed that infinitely often there are k -tuples where all the functions L_i have individually at most $C \log k$ prime factors, for an explicit constant C .

2010 *Mathematics Subject Classification*. 11N05, 11N35, 11N36.
Supported by EPSRC Doctoral Training Grant EP/P505216/1.

A different approach was taken by Goldston, Pintz and Yıldırım [4] in their work on small gaps between primes. Under the Elliot-Halberstam conjecture, they showed that there are infinitely many n for which at least two of $n, n + 4, n + 6, n + 10, n + 12, n + 16$ are prime. Thus there must be at least one specific 2-tuple where both functions are prime infinitely often if the Elliot-Halberstam conjecture holds.

2. STATEMENT OF RESULTS

Our main result is

Theorem 2.1. *Given a set of k admissible linear functions, for infinitely many $n \in \mathbb{N}$ the product $\Pi(n)$ has at most r_k prime factors, where r_k is given in Table 1 below.*

TABLE 1. Bounds for $\Omega(\Pi(n))$

k	3	4	5	6	7	8	9	10
r_k	8	11	15	18	22	26	30	34

Theorem 2.1 improves the previous best known bounds for $k \geq 4$, which were obtained by Diamond and Halberstam [2] for $4 \leq k \leq 6$ and by Ho and Tsang [6] for $7 \leq k \leq 10$. We fall just short of proving $r_k \leq 7$ for $k = 3$, and so fail to improve upon a result of Porter [9]. This comparison is shown in Table 2. We prove these results using a sieve

TABLE 2. Bounds for $\Omega(\Pi(n))$

k	3	4	5	6	7	8	9	10
Previous best bound	8	12	16	20	24	28	33	38
New bound	8	11	15	18	22	26	30	34

which is a combination of a weighted sieve similar to Selberg's $\Lambda^2 \Lambda^-$ sieve (see [10]), and the Graham-Goldston-Pintz-Yıldırım sieve (see [3]) used to count numbers with a specific number of prime factors.

We note that for k large our method only improves lower order terms, and so we do not improve the asymptotic bound $r_k \sim k \log k$.

In a forthcoming paper [7], we will also improve the bound when $k = 3$, using an argument based on the Diamon-Halberstam-Richert sieve rather than Selberg's sieve.

3. KEY IDEAS

We wish to show that for any sufficiently large N we have

$$(3.1) \quad \sum_{N \leq n \leq 2N} (c - \Omega(\Pi(n))) \left(\sum_{d|\Pi(n)} \lambda_d \right)^2 > 0$$

for some real numbers λ_d and some constant integer $c > 0$. From this it is clear that there must be some $n \in [N, 2N]$ such that $\Omega(\Pi(n)) \leq c$. Since this is true for all sufficiently large N , it follows that there are infinitely many integers n such that $\Omega(\Pi(n)) \leq c$.

The work of Heath-Brown [5] and Ho and Tsang [6] considered a similar sum, but used the divisor function $d(\Pi(n))$ instead of the number-of-prime-factors function Ω . Using the divisor function has the advantage that there are stronger level-of-distribution results available, but we find that this is outweighed by the fact that the Ω function is relatively much smaller than the divisor function on numbers with many prime factors.

The Ω function has Bombieri-Vinogradov style equidistribution results (as shown by Motohashi [8]), and so we would expect we should be able to estimate the above sum directly, in a method similar to Heath-Brown [5] or Selberg [10] when they considered the divisor function instead. We encounter some technical difficulties when attempting to translate this argument, however.

Instead we express $\Omega(n)$ as a weighted sum over small prime factors (as in the weighted sieve method of Diamond and Halberstam [2]) and a remaining positive contribution which we split up depending on the number of prime factors of each of the $L_j(n)$.

Diamond and Halberstam used a weighted sieve. The method relied on the fact for n square-free we have the inequality

$$(3.2) \quad \Omega(n) \leq \sum_{\substack{p|n \\ p \leq y}} \left(1 - \frac{\log p}{\log y}\right) + \frac{\log n}{\log y}.$$

We note that this inequality is strict if n has a prime factor which is larger than y . This results in a loss in the argument which has a noticeable effect when we apply this to k -tuples when k is small. Assuming that $y \geq n^{1/2}$ and n square-free we can write instead an equality

$$(3.3) \quad \Omega(n) = \sum_{\substack{p|n \\ p \leq y}} \left(1 - \frac{\log p}{\log y}\right) + \frac{\log n}{\log y} + \sum_{r=1}^{\infty} \chi_r(n),$$

where

$$(3.4) \quad \chi_r(n) = \begin{cases} 1 - \frac{\log p_r}{\log y}, & n = p_1 \dots p_r \text{ with } p_1 \leq p_2 \leq \dots \leq p_{r-1} \text{ and } y < p_r, \\ 0, & \text{otherwise,} \\ -\left(\frac{\log n}{\log y} - 1 - \sum_{i=1}^{r-1} \frac{\log p_i}{\log y}\right), & n = p_1 \dots p_r \text{ with } p_1 \leq p_2 \leq \dots \leq p_r \\ & \text{and } y < p_r, \\ 0, & \text{otherwise.} \end{cases}$$

For fixed r we can evaluate Selberg-type weighted sums over $\chi_r(L_i(n))$ using the method of Graham, Goldston, Pintz and Yıldırım in [3] as an extension of the original GPY method. We note that the contribution from $\chi_r(n)$ is always negative, so we can obtain a lower bound by simply omitting terms when $r > h$ for some constant h . The contribution of the χ_r terms decreases quickly with r , and so we in practice only need to calculate the contribution when r is small (in this paper we only consider the contributions of χ_r when $r \leq 4$). This is the key difference in our approach to previous methods, and allows us to obtain the improvements given by Theorem 2.1.

4. INITIAL CONSIDERATIONS

We adopt similar notation to that of Graham, Goldston, Pintz and Yıldırım in [3].

Let $\mathcal{L} = \{L_1, L_2, \dots, L_k\}$ be an admissible k -tuple of linear functions. We define

$$(4.1) \quad \Pi(n) = \prod_{i=1}^k L_i(n) = (a_1n + b_1) \dots (a_kn + b_k),$$

$$(4.2) \quad \nu_p(\mathcal{L}) = \#\{1 \leq n \leq p : \Pi(n) \equiv 0 \pmod{p}\}.$$

We note that admissibility is equivalent to the condition

$$(4.3) \quad \nu_p(\mathcal{L}) < p \quad \text{for all primes } p.$$

We also see that $\nu_p(\mathcal{L}) \leq k$ for all primes p , and so the above condition holds automatically for $p > k$.

For technical reasons we adopt a normalisation of our linear functions, as done originally by Heath-Brown in [5]. Since we are only interested in the showing any admissible k -tuple has at most r_k prime factors infinitely often (for some explicit r_k), by considering the functions $L_i(An + B)$ for suitably chosen constants A and B , we may assume without loss of generality that our functions satisfy the following hypothesis.

Hypothesis 1. $\mathcal{L} = \{L_1, \dots, L_k\}$ is an admissible k -tuple of linear functions. The functions $L_i(n) = a_i n + b_i$ ($1 \leq i \leq k$) are distinct with $a_i > 0$. Each of the coefficients a_i is composed of the same primes, none of which divides the b_j . If $i \neq j$, then any prime factor of $a_i b_j - a_j b_i$ divides each of the a_i .

For a set of linear functions satisfying Hypothesis 1 we define

$$(4.4) \quad A = \prod_{i=1}^k a_i.$$

We note that in this case

$$(4.5) \quad \nu_p(\mathcal{L}) = \begin{cases} 0, & p|A, \\ k, & p \nmid A. \end{cases}$$

We also define the *singular series* $\mathfrak{S}(\mathcal{L})$ of \mathcal{L} when \mathcal{L} satisfies Hypothesis 1.

$$(4.6) \quad \mathfrak{S}(\mathcal{L}) = \prod_{p|A} \left(1 - \frac{1}{p}\right)^{-k} \prod_{p \nmid A} \left(1 - \frac{k}{p}\right) \left(1 - \frac{1}{p}\right)^{-k}.$$

We note that $\mathfrak{S}(\mathcal{L})$ is positive.

As is common with the Selberg sieve, for some parameter R_2 we impose the condition

$$(4.7) \quad \lambda_d = 0 \quad \text{if } d \geq R_2 \text{ or } d \text{ not square-free or } (d, A) \neq 1.$$

We wish to choose the λ_d to maximize the sum (3.1), but this will be difficult to do optimally. We proceed by reparameterising the form in λ_d into new variables y_r and y_r^* which

will almost diagonalise it. We define

$$(4.8) \quad y_r = \mu(r) f_1(r) \sum_d' \frac{\lambda_{dr}}{f(dr)},$$

$$(4.9) \quad y_r^* = \mu(r) f_1^*(r) \sum_d' \frac{\lambda_{dr}}{f^*(dr)},$$

where here and from now on, the $'$ by the summation indicates that the sum is over all values of the indices which are square-free and coprime to A . For square-free d coprime to A , the functions f , f_1 , f^* and f_1^* are defined by

$$(4.10) \quad f(d) = \prod_{p|d} \frac{p}{k},$$

$$(4.11) \quad f_1(d) = (f * \mu)(d) = \prod_{p|d} \frac{p-k}{k},$$

$$(4.12) \quad f^*(d) = \prod_{p|d} \frac{p-1}{k-1},$$

$$(4.13) \quad f_1^*(d) = (f^* * \mu)(d) = \prod_{p|d} \frac{p-k}{k-1}.$$

We note that by Möbius inversion we have

$$(4.14) \quad \lambda_d = \mu(d) f(d) \sum_r' \frac{y_{rd}}{f_1(rd)}.$$

Thus the λ_d (and hence also the y_r^*) are defined uniquely by a choice of the y_r . The conditions (4.7) will be satisfied if the same conditions apply to the y_r .

For some polynomial P (to be determined later), we choose

$$(4.15) \quad y_r = \begin{cases} \mu^2(r) \mathfrak{S}(\mathcal{L}) P\left(\frac{\log R_2/r}{\log R_2}\right), & \text{if } r \leq R_2 \text{ and } (r, A) = 1, \\ 0, & \text{otherwise.} \end{cases}$$

We now turn our attention to the proof of the theorem.

5. PROOF OF THEOREM

We consider the sum

$$(5.1) \quad S = S(\nu; N, R_1, R_2, \mathcal{L}) = \sum_{N \leq n \leq 2N} w(n) \Lambda^2(n),$$

where

$$(5.2) \quad w(n) = \nu - \sum_{p|\Pi(n)} \left(1 - \frac{\log p}{\log R_1}\right),$$

$$(5.3) \quad \Lambda^2(n) = \left(\sum_{\substack{d|\Pi(n) \\ d \leq R_2}} \lambda_d \right)^2.$$

We note that if $\Pi(n)$ is square-free then

$$(5.4) \quad w(n) = \nu - \Omega(\Pi(n)) + \frac{\log \Pi(n)}{\log R_1}.$$

We see that for $n \in [N, 2N]$ and some fixed $h \in \mathbb{Z}_{>0}$ we have

$$\begin{aligned}
 w(n) &= \nu - \sum_{j=1}^k \sum_{p|L_j(n)} \left(1 - \frac{\log p}{\log R_1}\right) \\
 &\geq \nu - \sum_{j=1}^k \sum_{\substack{p|L_j(n) \\ p \leq R_1 \text{ or } \Omega(L_j(n)) \leq h}} \left(1 - \frac{\log p}{\log R_1}\right) \\
 (5.5) \quad &\geq \nu - \sum_{j=1}^k \sum_{\substack{p|L_j(n) \\ p \leq R_1}} \left(1 - \frac{\log p}{\log R_1}\right) + \sum_{j=1}^k \sum_{r=1}^h \chi_r(L_j(n)),
 \end{aligned}$$

where

$$(5.6) \quad \chi_r(n) = \begin{cases} \frac{\log N}{\log R_1} - 1 - \sum_{i=1}^{r-1} \frac{\log p_i}{\log R_1}, & \text{if } n = p_1 \dots p_r \text{ with} \\ & n^\epsilon < p_1 < \dots < p_{r-1} \leq n^{\log R_1 / \log N} < p_r \\ 0, & \text{otherwise.} \end{cases}$$

Thus

$$\begin{aligned}
 \sum_{\substack{N \leq n \leq 2N \\ \Pi(n) \text{ square-free}}} \left(\nu - \Omega(\Pi(n)) + \frac{\log \Pi(n)}{\log R_1} \right) \Lambda^2(n) &= S - S' \\
 (5.7) \quad &\geq \nu S_0 - S' - T_0 + \sum_{j=1}^k \sum_{r=1}^h T_{r,j},
 \end{aligned}$$

where

$$(5.8) \quad S_0 = \sum_{N \leq n \leq 2N} \Lambda^2(n),$$

$$(5.9) \quad S' = \sum_{\substack{N \leq n \leq 2N \\ \Pi(n) \text{ not square-free}}} w(n) \Lambda^2(n),$$

$$(5.10) \quad T_0 = \sum_{N \leq n \leq 2N} \sum_{\substack{p|\Pi(n) \\ p \leq R_1}} \left(1 - \frac{\log p}{\log R_1}\right) \Lambda^2(n),$$

$$(5.11) \quad T_{r,j} = \sum_{N \leq n \leq 2N} \chi_r(L_j(n)) \Lambda^2(n).$$

We can evaluate S_0 , S' , T_0 and T_i using weighted forms of the Selberg sieve. We state the results here and prove them in the following sections. To ease notation we now fix as constants

$$(5.12) \quad r_1 = \frac{\log R_1}{\log N}, \quad r_2 = \frac{\log R_2}{\log N}.$$

We view r_1 , r_2 , k , A and our polynomial P as fixed, and so any constants implied by the use of O or \ll notation may depend on these quantities without explicit reference.

Proposition 5.1. *Let \mathcal{L} satisfy Hypothesis 1. Let $W_0 : [0, r_1/r_2] \rightarrow \mathbb{R}_{\geq 0}$ be a piecewise smooth non-negative function. Let λ_d, y_d be as given in (4.14) and (4.15). Assume that $r_1 \geq r_2$. Then there exists a constant C such that if $R_1 R_2^2 \leq N(\log N)^{-C}$ then we have*

$$\sum_{N \leq n \leq 2N} \left(\sum_{\substack{p|\Pi(n) \\ p \leq R_1}} W_0 \left(\frac{\log p}{\log R_2} \right) \right) \left(\sum_{\substack{d|\Pi(n) \\ d \leq R_2}} \lambda_d \right)^2 = \frac{\mathfrak{S}(\mathcal{L}) N (\log R_2)^k}{(k-1)!} J_0 \\ + O_{W_0} \left(N (\log N)^{k-1} (\log \log N)^2 \right),$$

where

$$J_0 = J_{01} + J_{02} + J_{03}, \\ J_{01} = k \int_0^1 \frac{W_0(y)}{y} \int_0^{1-y} (P(1-x) - P(1-x-y))^2 x^{k-1} dx dy, \\ J_{02} = k \int_0^1 \frac{W_0(y)}{y} \int_{1-y}^1 P(1-x)^2 x^{k-1} dx dy, \\ J_{03} = k \int_1^{r_1/r_2} \frac{W_0(y)}{y} \int_0^1 P(1-x)^2 x^{k-1} dx dy.$$

Proposition 5.2. *Given $\epsilon > 0$ and $r \in \mathbb{Z}_{>0}$, let*

$$\mathcal{A}_r := \left\{ x \in [0, 1]^{r-1} : \epsilon < x_1 < \dots < x_{r-1}, \sum_{i=1}^{r-1} x_i < \min(1 - r_2, 1 - x_{r-1}) \right\}.$$

Let $W_r : [0, 1]^{r-1} \rightarrow \mathbb{R}_{\geq 0}$ be a piecewise smooth function supported on \mathcal{A}_r such that

$$\frac{\partial}{\partial x_j} W_r(x) \ll W_r(x) \quad \text{uniformly for } x \in \mathcal{A}_r.$$

Let

$$\beta_r(n) = \begin{cases} W_r \left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_{r-1}}{\log n} \right), & n = p_1 p_2 \dots p_r \text{ with } p_1 < \dots < p_r, \\ 0, & \text{otherwise,} \end{cases}$$

Then there is a constant C such that if $R_2^2 \leq N^{1/2} (\log N)^{-C}$, we have

$$\sum_{N \leq n \leq 2N} \beta_r(L_j(n)) \left(\sum_{\substack{d|\Pi(n) \\ d \leq R_2}} \lambda_d \right)^2 = \frac{\mathfrak{S}(\mathcal{L}) N (\log R_2)^{k+1}}{(k-2)! (\log N)} J_r + O_{W_r} \left(N (\log \log N)^r (\log N)^{k-1} \right),$$

where

$$J_r = \int_{(x_1, \dots, x_{r-1}) \in \mathcal{A}_r} \frac{W_r(x_1, \dots, x_{r-1}) I_1(r_2^{-1} x_1, \dots, r_2^{-1} x_{r-1})}{\left(\prod_{i=1}^{r-1} x_i \right) \left(1 - \sum_{i=1}^{r-1} x_i \right)} dx_1 \dots dx_{r-1}, \\ I_1 = \int_0^1 \left(\sum_{J \subset \{1, \dots, r-1\}} (-1)^{|J|} \tilde{P}^+(1-t - \sum_{i \in J} x_i) \right)^2 t^{k-2} dt, \\ \tilde{P}^+(x) = \begin{cases} \int_0^x P(t) dt, & x \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

Proposition 5.3. *There exists a constant C such that if $R_2^2 \leq N^{1/2}(\log N)^{-C}$ then*

$$\sum_{\substack{N \leq n \leq 2N \\ \Pi(n) \text{ not square-free}}} \left(\sum_{\substack{d|\Pi(n) \\ d \leq R_2}} \lambda_d \right)^2 \ll N(\log N)^{k-1} \log \log N.$$

We also quote a result [3][Theorem 7] which is based on the original result of Goldston, Pintz and Yıldırım in [4].

Proposition 5.4. *There is a constant C such that if $R_2^2 \leq N(\log N)^{-C}$, we have*

$$\sum_{N \leq n \leq 2N} \left(\sum_{\substack{d|\Pi(n) \\ d \leq R_2}} \lambda_d \right)^2 = \frac{\mathfrak{S}(\mathcal{L})N(\log R_2)^k}{(k-1)!} J + O(N(\log N)^{k-1})$$

where

$$J = \int_0^1 P(1-t)^2 t^{k-1} dt.$$

Using Propositions 5.1, 5.2, 5.4 and 5.3 we can now bound our sum S in terms of the integers k and h and the polynomial P . For some $\epsilon > 0$ we choose

$$(5.13) \quad r_1 = \frac{1}{2} + \epsilon, \quad r_2 = \frac{1}{4} - \epsilon,$$

so that the conditions of all the propositions are satisfied.

Proposition 5.4 gives the size of S_0 immediately.

Using Proposition 5.3 we have

$$\begin{aligned} S' &= \sum_{\substack{N \leq n \leq 2N \\ \Pi(n) \text{ not square-free}}} w(n) \Lambda^2(n) \\ &\leq \sum_{\substack{N \leq n \leq 2N \\ \Pi(n) \text{ not square-free}}} \left(\nu + \frac{\log \Pi(n)}{\log R_1} \right) \Lambda^2(n) \\ &\leq \sum_{\substack{N \leq n \leq 2N \\ \Pi(n) \text{ not square-free}}} \left(\nu + \frac{k + \epsilon}{r_1} \right) \Lambda^2(n) \\ (5.14) \quad &\ll N(\log N)^{k-1} \log \log N. \end{aligned}$$

To estimate T_0 and the $T_{r,j}$ we choose

$$(5.15) \quad W_0(x) = 1 - \frac{r_2}{r_1} x,$$

$$(5.16) \quad W_j(x_1, \dots, x_{j-1}) = \begin{cases} \frac{1}{r_1} - 1 - \frac{1}{r_1} \sum_{i=1}^{j-1} x_i, & \epsilon < x_1 < \dots < x_{j-1} \\ & \text{and } \sum_{i=1}^{r-1} x_i < 1 - r_1 \\ 0, & \text{otherwise,} \end{cases}$$

which satisfy the conditions of Propositions 5.1 and 5.2 respectively.

By Proposition 5.1 we have

$$\begin{aligned}
 T_0 &= \sum_{N \leq n \leq 2N} \left(\sum_{\substack{p | \Pi(n) \\ p \leq R_1}} W_0 \left(\frac{\log p}{\log R_2} \right) \right) \left(\sum_{\substack{d | \Pi(n) \\ d \leq R_2}} \lambda_d \right)^2 \\
 (5.17) \quad &= \frac{\mathfrak{S}(\mathcal{L}) N (\log R_2)^k}{(k-1)!} J_0 + O(N (\log N)^{k-1} \log \log N)
 \end{aligned}$$

where

$$(5.18) \quad J_0 = J_{01} + J_{02} + J_{03},$$

$$(5.19) \quad J_{01} = k \int_0^1 \frac{r_1 - r_2 y}{r_1 y} \int_0^{1-y} (P(1-x) - P(1-x-y))^2 x^{k-1} dx dy,$$

$$(5.20) \quad J_{02} = k \int_0^1 \frac{r_1 - r_2 y}{r_1 y} \int_{1-y}^1 P(1-x)^2 x^{k-1} dx dy,$$

$$(5.21) \quad J_{03} = k \int_1^{r_1/r_2} \frac{r_1 - r_2 y}{r_1 y} \int_0^1 P(1-x)^2 x^{k-1} dx dy.$$

By Proposition 5.2 we have

$$\begin{aligned}
 T_{r,j} &= \sum_{N \leq n \leq 2N} \chi_r(L_j(n)) \Lambda^2(n) \\
 &= \sum_{N \leq n \leq 2N} \beta_r(L_j(n)) \Lambda^2(n) \\
 (5.22) \quad &= \frac{\mathfrak{S}(\mathcal{L}) N (\log R_2)^{k+1}}{(k-2)! (\log N)} J_r + O_r(N (\log \log N)^{r+1} (\log N)^{k-1}),
 \end{aligned}$$

where

$$(5.23) \quad \beta_r(n) = \begin{cases} W_r \left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_{r-1}}{\log n} \right), & n = p_1 p_2 \dots p_r \text{ with } p_1 < \dots < p_r, \\ 0, & \text{otherwise,} \end{cases}$$

$$(5.24) \quad J_r = \int_{(x_1, \dots, x_{r-1}) \in \mathcal{A}_r} \frac{W_r(x_1, \dots, x_{r-1}) I_1(r_2^{-1} x_1, \dots, r_2^{-1} x_{r-1})}{\left(\prod_{i=1}^{r-1} x_i \right) \left(1 - \sum_{i=1}^{r-1} x_i \right)} dx_1 \dots dx_{r-1}.$$

Therefore we see that

$$\begin{aligned}
 (5.25) \quad \nu S_0 - S' + T_0 + \sum_{j=1}^k \sum_{r=1}^h T_{r,j} &= \frac{N \mathfrak{S}(\mathcal{L}) (\log R_2)^k}{(k-1)!} \left(\nu J - J_0 + r_2 k(k-1) \left(\sum_{r=1}^h J_r \right) \right) \\
 &\quad + O \left(\frac{N (\log N)^k}{\log \log N} \right).
 \end{aligned}$$

Therefore we put

$$(5.26) \quad \nu = \frac{J_0 - r_2 k(k-1) (\sum_{r=1}^h J_r)}{J} + \epsilon.$$

We then see that for any N sufficiently large we have

$$(5.27) \quad \nu S_0 - S' - T_0 + \sum_{j=1}^k \sum_{r=1}^h T_{r,j} > 0.$$

Thus we have

$$(5.28) \quad \Omega(\Pi(n)) \leq \left\lfloor \frac{J_0 - r_2 k(k-1)(\sum_{r=1}^h J_r)}{J} + \frac{k}{r_1} + 2\epsilon \right\rfloor$$

infinitely often.

With these fixed, given k, h and a polynomial P we obtain a bound on $\Omega(\Pi(n))$. To make calculations feasible we choose $h = 3$ (except we take $h = 4$ when $k = 10$). Numerical experiments indicate that the bounds of Theorem 1 cannot be improved by increasing h except possibly when $k = 5$.

We can now explicitly write down the integrals J_1 , J_2 and J_3 , splitting the integral up depending on whether \tilde{P}^+ is positive or not. We put

$$(5.29) \quad \tilde{P}(x) = \int_0^x P(t)dt.$$

Then we have that

$$(5.30) \quad J_1 = \left(\frac{1-r_1}{r_1} \right) \int_0^1 \tilde{P}(1-x)^2 x^{k-2} dx + O(\epsilon).$$

Similarly

$$(5.31) \quad J_2 = J_{21} + J_{22} + J_{23} + O(\epsilon),$$

where

$$(5.32) \quad J_{21} = \int_0^1 \frac{1-r_1-r_2y}{r_1y(1-r_2y)} \int_0^{1-y} (\tilde{P}(1-x) - \tilde{P}(1-x-y))^2 x^{k-2} dx dy,$$

$$(5.33) \quad J_{22} = \int_0^1 \frac{1-r_1-r_2y}{r_1y(1-r_2y)} \int_{1-y}^1 \tilde{P}(1-x)^2 x^{k-2} dx dy,$$

$$(5.34) \quad J_{23} = \int_1^{(1-r_1)/r_2} \frac{1-r_1-r_2y}{r_1y(1-r_2y)} \int_0^1 \tilde{P}(1-x)^2 x^{k-2} dx dy.$$

Finally

$$(5.35) \quad J_3 = J_{31} + J_{32} + J_{33} + J_{34} + J_{35} + J_{36} + J_{37} + J_{38} + O(\epsilon),$$

where

$$(5.36) \quad J_{31} = \int_1^{(1-r_1)/2r_2} \int_y^{(1-r_1)/r_2-y} \frac{1-r_1-r_2(y+z)}{r_1yz(1-r_2(y+z))} \int_0^1 \tilde{P}(1-x)^2 x^{k-2} dx dz dy,$$

$$(5.37) \quad J_{32} = \int_0^1 \int_y^{(1-r_1)/r_2-y} \frac{1-r_1-r_2(y+z)}{r_1yz(1-r_2(y+z))} \int_{1-y}^1 \tilde{P}(1-x)^2 x^{k-2} dx dz dy,$$

$$(5.38) \quad J_{33} = \int_0^1 \int_1^{(1-r_1)/r_2-y} \frac{1-r_1-r_2(y+z)}{r_1yz(1-r_2(y+z))} \int_0^{1-y} (\tilde{P}(1-x) - \tilde{P}(1-x-y))^2 x^{k-2} dx dz dy,$$

$$(5.39) \quad J_{34} = \int_0^1 \int_y^1 \frac{1-r_1-r_2(y+z)}{r_1yz(1-r_2(y+z))} \int_{1-z}^{1-y} (\tilde{P}(1-x) - \tilde{P}(1-x-y))^2 x^{k-2} dx dz dy,$$

$$(5.40) \quad J_{35} = \int_{1/2}^1 \int_y^1 \frac{1 - r_1 - r_2(y+z)}{r_1 y z (1 - r_2(y+z))} \int_0^{1-z} \left(\tilde{P}(1-x) - \tilde{P}(1-x-y) - \tilde{P}(1-x-z) \right)^2 x^{k-2} dx dz dy,$$

$$(5.41) \quad J_{36} = \int_0^{1/2} \int_{1-y}^1 \frac{1 - r_1 - r_2(y+z)}{r_1 y z (1 - r_2(y+z))} \int_0^{1-z} \left(\tilde{P}(1-x) - \tilde{P}(1-x-y) - \tilde{P}(1-x-z) \right)^2 x^{k-2} dx dz dy,$$

$$(5.42) \quad J_{37} = \int_0^{1/2} \int_y^{1-y} \frac{1 - r_1 - r_2(y+z)}{r_1 y z (1 - r_2(y+z))} \int_{1-y-z}^{1-z} \left(\tilde{P}(1-x) - \tilde{P}(1-x-y) - \tilde{P}(1-x-z) \right)^2 x^{k-2} dx dz dy,$$

$$(5.43) \quad J_{38} = \int_0^{1/2} \int_y^{1-y} \frac{1 - r_1 - r_2(y+z)}{r_1 y z (1 - r_2(y+z))} \int_0^{1-y-z} \left(\tilde{P}(1-x) - \tilde{P}(1-x-y) - \tilde{P}(1-x-z) + \tilde{P}(1-x-y-z) \right)^2 x^{k-2} dx dz dy.$$

We now have explicit representations of J , J_0 , J_1 , J_2 and J_3 . We can calculate these by numerical integration given k and a polynomial P .

Table 3 gives close to optimal polynomials for $3 \leq k \leq 10$ and the corresponding bounds obtained if we take ϵ sufficiently small. These give the results claimed in Theorem 2.1 except for $k = 10$.

TABLE 3. Bounds for $\Omega(\Pi(n))$

k	Bound on $\Omega(\Pi(n))$	Polynomial $P(x)$
3	8.220...	$1 + 14x$
4	11.653...	$1 + 22x$
5	15.306...	$1 + 33x$
6	18.936...	$1 + 10x + 40x^2$
7	22.834...	$1 + 10x + 60x^2$
8	26.860...	$1 + 10x + 80x^2$
9	30.942...	$1 + 30x + 300x^3$
10	35.158...	$1 + 35x - 10x^2 + 400x^3$

For $k = 10$ we find an improvement if we also include the contribution when one of the $L_i(n)$ has 4 prime factors (we omit the explicit integrals here). In this case we choose the polynomial

$$(5.44) \quad P(x) = 1 + 10x + 150x^2.$$

This gives us the bound 34.77... and so 10-tuples infinitely often have at most 34 prime factors, verifying Theorem 1.

6. THE QUANTITIES T_δ AND T_δ^*

Before proving the propositions, we first establish some results about the quantities

$$(6.1) \quad T_\delta = \sum'_{d,e} \frac{\lambda_d \lambda_e}{f([d, e, \delta]/\delta)},$$

$$(6.2) \quad T_\delta^* = \sum'_{d,e} \frac{\lambda_d \lambda_e}{f^*([d, e, \delta]/\delta)}.$$

Most of these results already exist in some form in the literature. These results will underlie the proof of the propositions. We note that in [3] Graham, Goldston, Pintz and Yıldırım used slightly different notation (our quantity T_δ^* is labelled T_δ).

We first put T_δ and T_δ^* into an almost-diagonalised form.

Lemma 6.1. *We have*

$$T_\delta = \sum'_{\substack{a \\ (a, \delta)=1}} \frac{\mu^2(a)}{f_1(a)} \left(\sum_{s|\delta} \mu(s) y_{as} \right)^2,$$

$$T_\delta^* = \sum'_{\substack{a \\ (a, \delta)=1}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) y_{as}^* \right)^2,$$

where

$$y_a^* = \frac{\mu^2(a)a}{\phi(a)} \sum_m' \frac{y_{ma}}{\phi(m)}.$$

Proof. The result for T_δ is shown, for example, in [10][Page 85]. The result for T_δ^* is proven in [3][Lemma 6]. \square

We now again quote a Lemma from [3], which expresses the y_a^* in terms of the polynomial P which we used to define the variables y_a .

Lemma 6.2. *Let*

$$y_a = \begin{cases} \mu^2(a) \mathfrak{S}(\mathcal{L}) P\left(\frac{\log R_2/a}{\log R_2}\right), & \text{if } 0 \leq a < R_2 \text{ and } (a, A) = 1 \\ 0, & \text{otherwise} \end{cases}.$$

Then we have for $(a, A) = 1$ and $a < R_2$ that

$$y_a^* = \mu^2(a) \frac{\phi(A)}{A} \mathfrak{S}(\mathcal{L}) (\log R_2) \tilde{P}\left(\frac{\log R_2/a}{\log R_2}\right) + O(\log \log R_2),$$

where

$$\tilde{P}(x) = \int_0^x P(t) dt.$$

If $(a, A) \neq 1$ or $a \geq R_2$ then we have

$$y_a^* = 0.$$

Proof. This is proven in [3][Lemma 7]. \square

We will repeatedly use the following result.

Lemma 6.3. *For $u \geq 1$ we have*

$$\begin{aligned} \sum'_{a < u} \frac{\mu^2(a)}{f_1(a)} &= \frac{A}{\phi(A)} \frac{(\log u)^k}{\mathfrak{S}(\mathcal{L})k!} + O((\log 2u)^{k-1}), \\ \sum'_{a < u} \frac{\mu^2(a)}{f_1^*(a)} &= \frac{A}{\phi(A)} \frac{(\log u)^{k-1}}{\mathfrak{S}(\mathcal{L})(k-1)!} + O((\log 2u)^{k-2}). \end{aligned}$$

Proof. This follows, for example, from [3][Lemma 3]. \square

In order to estimate the terms T_δ^* we wish to remove the condition $(a, \delta) = 1$ in the summation over a , and remove the constraint caused by y_a and y_a^* only being supported on square-free a . We let

$$(6.3) \quad P_a = \begin{cases} \mathfrak{S}(\mathcal{L})P\left(\frac{\log R_2/a}{\log R_2}\right), & \text{if } 0 \leq a < R_2 \\ 0, & \text{otherwise,} \end{cases}$$

$$(6.4) \quad P_a^* = \begin{cases} \frac{\phi(A)}{A} \mathfrak{S}(\mathcal{L})(\log R_2) \tilde{P}\left(\frac{\log R_2/a}{\log R_2}\right), & \text{if } 0 \leq a < R_2 \\ 0, & \text{otherwise,} \end{cases}$$

so that these are equal to y_a and $y_a^* + O(\log \log R_2)$ respectively when a is square-free and coprime to A .

Lemma 6.4. *Let $(\delta, A) = 1$. Then we have*

$$\begin{aligned} T_\delta &= \sum'_a \frac{\mu^2(a)}{f_1(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* \right)^2 + O\left(d(\delta)^2 (\log R_2)^{k-1} \log \log R_2\right), \\ T_\delta^* &= \sum'_a \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* \right)^2 + O\left(d(\delta)^2 (\log R_2)^k \log \log R_2\right). \end{aligned}$$

Proof. We only prove the result for the T_δ^* here, the result for the T_δ follows from a completely analogous argument. We see that since $P_a^* \ll \log R_2$ we have

$$\begin{aligned} T_\delta^* &= \sum'_{\substack{a \\ (a, \delta)=1}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* + O(\log \log R_2) \right)^2 \\ (6.5) \quad &= \sum'_{\substack{a \\ (a, \delta)=1}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* \right)^2 + O\left(d(\delta)^2 (\log R_2) (\log \log R_2) \sum_{a < R_2} \frac{\mu^2(a)}{f_1^*(a)}\right). \end{aligned}$$

By Lemma 6.3 the error term above is $O(d(\delta)^2 (\log R_2)^k \log \log R_2)$.

We see that to prove the result it is sufficient to prove

$$(6.6) \quad \sum'_{\substack{a \\ (a, \delta) \neq 1}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* \right)^2 \ll (\log R_2)^k d(\delta)^2 (\log \log R_2).$$

Since all terms in the sum are non-negative, we have

$$(6.7) \quad \sum'_{\substack{a \\ (a, \delta) \neq 1}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* \right)^2 \leq \sum_{p|\delta} \sum'_{\substack{a \\ p|a}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* \right)^2.$$

We consider the inner sum. By the Cauchy-Schwarz inequality we have

$$(6.8) \quad \sum'_{\substack{a \\ p|a}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* \right)^2 = \sum'_{\substack{a \\ p|a}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta/p} \mu(s) (P_{as}^* - P_{asp}^*) \right)^2 \\ \ll d(\delta) \sum_{s|\delta/p} \sum'_{\substack{a \\ p|a}} \frac{\mu^2(a)}{f_1^*(a)} (P_{as}^* - P_{asp}^*)^2.$$

We split the summation over a depending on whether the P_{as}^* and P_{asp}^* terms vanish (since $P_b^* = 0$ for $b \geq R_2$).

$$(6.9) \quad \sum'_{\substack{a \\ p|a}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* \right)^2 \ll d(\delta) \sum_{s|\delta/p} \sum'_{a' < R_2/sp^2} \frac{\mu^2(a'p)}{f_1^*(a'p)} (P_{a'ps}^* - P_{a'sp^2}^*)^2 \\ + d(\delta) \sum_{s|\delta/p} \sum'_{R_2/sp^2 \leq a' < R_2/sp} \frac{\mu^2(a'p)}{f_1^*(a'p)} (P_{a'ps}^*)^2.$$

We substitute in the value of P^* .

$$(6.10) \quad \frac{1}{d(\delta)} \sum'_{\substack{a \\ p|a}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|q} \mu(s) P_{as}^* \right)^2 \\ \ll (\log R_2)^2 \sum_{s|\delta/p} \sum'_{a' < R_2/sp^2} \frac{\mu^2(a'p)}{f_1^*(a'p)} \left(\tilde{P} \left(1 - \frac{\log a'ps}{\log R_2} \right) - \tilde{P} \left(1 - \frac{\log a'sp^2}{\log R_2} \right) \right)^2 \\ + (\log R_2)^2 \sum_{s|\delta/p} \sum'_{R_2/sp^2 \leq a' < R_2/sp} \frac{\mu^2(a'p)}{f_1^*(a'p)} \tilde{P} \left(1 - \frac{\log a'ps}{\log R_2} \right)^2.$$

In the first sum above both the arguments of the polynomials differ by $\log p / \log R_2$. Since they are fixed polynomials, the derivative of the polynomial is $\ll 1$ and so the difference is $\ll \log p / \log R_2$. In the second sum we just use the trivial bound $\tilde{P}(x) \ll 1$.

This gives

$$(6.11) \quad \frac{1}{d(\delta)} \sum'_{\substack{a \\ p|a}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|q} \mu(s) P_{as}^* \right)^2 \ll \frac{(\log p)^2}{f_1^*(p)} \sum_{s|\delta/p} \sum'_{a < R_2/sp^2} \frac{\mu^2(a)}{f_1^*(a)} \\ + \frac{(\log R_2)^2}{f(p)} \sum_{s|\delta/p} \sum'_{R_2/sp^2 \leq a < R_2/sp} \frac{\mu^2(a)}{f_1^*(a)}.$$

Using Lemma 6.3 we see that the first sum is $\ll d(\delta)(\log p)^2(\log R_2)^{k-1}/f_1^*(p)$ and the second sum is $\ll d(\delta)(\log p)(\log R_2)^k/f_1^*(p)$ because of the range of summation over a . Thus

$$(6.12) \quad \sum'_{\substack{a \\ p|a}} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* \right)^2 \ll d(\delta)^2 \frac{\log p}{f_1^*(p)} (\log R_2)^k \ll d(\delta)^2 \frac{\log p}{p} (\log R_2)^k.$$

Summing over all $p|\delta$ gives the bound

$$(6.13) \quad d(\delta)^2 (\log R_2)^k \sum_{p|\delta} \frac{\log p}{p}.$$

Splitting the sum into a sum over $p \leq \log R_2$ and a sum over $p > \log R_2$ we get the bound

$$(6.14) \quad d(\delta)^2 (\log R_2)^k (\log \log R_2).$$

This gives (6.6), and hence the Lemma. \square

Essentially the same argument as above also yields a useful bound on the size of T_δ and T_δ^* .

Lemma 6.5. *Let $(\delta, A) = 1$. Then we have*

$$\begin{aligned} T_\delta &\ll \min_{p|\delta} (\log p) d(\delta)^2 (\log R_2)^{k-1}, \\ T_\delta^* &\ll \min_{p|\delta} (\log p) d(\delta)^2 (\log R_2)^k + d(\delta)^2 (\log R_2)^k \log \log R_2. \end{aligned}$$

Proof. For $p|\delta$ we have (using the fact all terms are non-negative)

$$\begin{aligned} T_\delta^* &= \sum_{\substack{a \\ (a, \delta)=1}}' \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* \right)^2 + O(d(\delta)^2 (\log R_2)^k \log \log R_2) \\ &\ll d(\delta) \sum_a' \frac{\mu^2(a)}{f_1^*(a)} \sum_{s|\delta/p} (P_{as}^* - P_{asp}^*)^2 + d(\delta)^2 (\log R_2)^k \log \log R_2 \\ &\ll d(\delta) (\log R_2)^2 \sum_{s|\delta/pa < R_2/sp} \sum_{sp \leq a < R_2/s}' \frac{\mu^2(a)}{f_1^*(a)} \left(\tilde{P} \left(1 - \frac{\log as}{\log R_2} \right) - \tilde{P} \left(1 - \frac{\log asp}{\log R_2} \right) \right)^2 \\ &\quad + d(\delta) (\log R_2)^2 \sum_{s|\delta/pR_2/sp \leq a < R_2/s}' \frac{\mu^2(a)}{f_1^*(a)} \tilde{P} \left(1 - \frac{\log as}{\log R_2} \right)^2 \\ (6.15) \quad &+ d(\delta)^2 (\log R_2)^k \log \log R_2. \end{aligned}$$

Noting the difference of the polynomials in the first sum is $\ll \log p / \log R_2$, and the polynomial in the second sum is $\ll 1$, we have

$$\begin{aligned} T_\delta^* &\ll d(\delta) (\log p)^2 \sum_{s|\delta/pa < R_2/sp} \sum_{sp \leq a < R_2/s}' \frac{\mu^2(a)}{f_1^*(a)} + d(\delta) (\log R_2)^2 \sum_{s|\delta/pR_2/sp \leq a < R_2/s}' \frac{\mu^2(a)}{f_1^*(a)} \\ (6.16) \quad &+ d(\delta)^2 (\log R_2)^k \log \log R_2. \end{aligned}$$

Appealing to Lemma 6.3 as in the previous lemma we obtain

$$(6.17) \quad T_\delta^* \ll d(\delta)^2 (\log p) (\log R_2)^k + d(\delta)^2 (\log R_2)^k \log \log R_2.$$

The result for T_δ follows by a completely analogous argument. In this case the first line holds without the $O(d(\delta)^2 (\log R_2)^k \log \log R_2)$ term, and so the final expression also holds without this term. \square

With these results we are able to get an integral expression for T_δ and T_δ^* when δ has a bounded number of prime factors.

Lemma 6.6. *Let $p_1, \dots, p_{r-1} \nmid A$ for some primes p_1, \dots, p_{r-1} . Then we have*

$$\begin{aligned} T_{p_1 \dots p_{r-1}} &= (\log R_2)^k \frac{\mathfrak{S}(\mathcal{L})}{(k-1)!} I_0 \left(\frac{\log p_1}{\log R_2}, \dots, \frac{\log p_{r-1}}{\log R_2} \right) \\ &\quad + O_r((\log R_2)^{k-1} \log \log R_2), \\ T_{p_1 \dots p_{r-1}}^* &= (\log R_2)^{k+1} \frac{\phi(A) \mathfrak{S}(\mathcal{L})}{A(k-2)!} I_1 \left(\frac{\log p_1}{\log R_2}, \dots, \frac{\log p_{r-1}}{\log R_2} \right) \\ &\quad + O_r((\log R_2)^k \log \log R_2). \end{aligned}$$

Here

$$\begin{aligned} I_0(x_1, \dots, x_{r-1}) &= \int_0^1 \left(\sum_{J \subset \{1, \dots, r-1\}} P^+ \left(1 - t - \sum_{j \in J} x_j \right) (-1)^{|J|} \right)^2 t^{k-1} dt, \\ I_1(x_1, \dots, x_{r-1}) &= \int_0^1 \left(\sum_{J \subset \{1, \dots, r-1\}} \tilde{P}^+ \left(1 - t - \sum_{j \in J} x_j \right) (-1)^{|J|} \right)^2 t^{k-2} dt, \\ P^+(x) &= \begin{cases} P(x), & x \geq 0, \\ 0, & \text{otherwise,} \end{cases} \\ \tilde{P}^+(x) &= \begin{cases} \int_0^x P(t) dt, & x \geq 0, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Proof. Let $\delta = p_1 \dots p_{r-1}$.

By Lemmas 6.1 and 6.4 we have that

$$(6.18) \quad T_\delta^* = \sum_a' \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) P_{as}^* \right)^2 + O_r((\log R_2)^k \log \log R_2).$$

We recall from (6.3) that for $a < R_2$ we have

$$(6.19) \quad P_a^* = \mu^2(a) \frac{A}{\phi(A)} (\log R_2) \mathfrak{S}(\mathcal{L}) \tilde{P}^+ \left(\frac{\log R_2/a}{\log R_2} \right).$$

Substituting this in above for $(\delta, A) = 1$ we obtain

$$\begin{aligned} (6.20) \quad T_\delta^* &= \frac{A^2}{\phi(A)^2} (\log R_2)^2 \mathfrak{S}(\mathcal{L})^2 \sum_{(a,A)=1} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|\delta} \mu(s) \tilde{P}^+ \left(\frac{\log R_2/as}{\log R_2} \right) \right)^2 \\ &\quad + O_r((\log R_2)^k \log \log R_2). \end{aligned}$$

We again use Lemma 6.3 which shows that

$$(6.21) \quad \sum_{\substack{a \leq R_2 \\ (a,A)=1}} \frac{\mu^2(a)}{f_1^*(a)} \ll (\log R_2)^{k-1}.$$

Thus

$$\begin{aligned} (6.22) \quad T_{p_1 \dots p_{r-1}}^* &= \frac{A^2}{\phi(A)^2} (\log R_2)^2 \mathfrak{S}(\mathcal{L})^2 \sum_{(a,A)=1} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|p_1 \dots p_{r-1}} \mu(s) \tilde{P}^+ \left(\frac{\log R_2/as}{\log R_2} \right) \right)^2 \\ &\quad + O_r((\log R_2)^k (\log \log R_2)). \end{aligned}$$

We also have

$$(6.23) \quad T_{p_1 \dots p_{r-1}} = \mathfrak{S}(\mathcal{L})^2 \sum_{(a,A)=1} \frac{\mu^2(a)}{f_1(a)} \left(\sum_{s|p_1 \dots p_{r-1}} \mu(s) P^+ \left(\frac{\log R_2/as}{\log R_2} \right) \right)^2 + O_r((\log R_2)^{k-1}).$$

We can now estimate the main term using [3][Lemma 4]. First we put

$$(6.24) \quad \gamma(p) = \begin{cases} k-1, & p \nmid A \\ 0, & \text{otherwise.} \end{cases}$$

$$g(d) = \prod_{p|d} \frac{\gamma(p)}{p - \gamma(p)},$$

$$F(t) = F_{x_1, \dots, x_{r-1}}(t) = \sum_{J \subset \{1, \dots, r-1\}} (-1)^{|J|} \tilde{P}^+ \left(t + \sum_{j \in J} x_j \right).$$

If we put $x_i = \log p_i / \log R_2$ for each $i \in \{1, \dots, r-1\}$ then we see that

$$(6.25) \quad \sum_{(a,A)=1} \frac{\mu^2(a)}{f_1^*(a)} \left(\sum_{s|p_1 \dots p_{r-1}} \mu(s) \tilde{P}^+ \left(\frac{\log R_2/as}{\log R_2} \right) \right)^2 = \sum_{d \leq R_2} \mu^2(d) g(d) F \left(\frac{\log R_2/d}{\log R_2} \right).$$

Since F is a continuous piecewise differentiable function we can apply [3][Lemma 4] which gives

$$(6.26) \quad \sum_{d \leq R_2} \mu^2(d) g(d) F \left(\frac{\log R_2/d}{\log R_2} \right) = \frac{A}{\phi(A)} \frac{(\log R_2)^{k-1}}{\mathfrak{S}(\mathcal{L})(k-2)!} \int_0^1 F(1-t) t^{k-2} dt + O((\log R_2)^{k-2}).$$

Similarly we follow the same procedure instead with

$$(6.27) \quad \gamma(p) = \begin{cases} k, & p \nmid A \\ 0, & \text{otherwise,} \end{cases}$$

$$G(t) = \sum_{J \subset \{1, \dots, r-1\}} (-1)^{|J|} P^+ \left(t + \sum_{j \in J} x_j \right).$$

This yields

$$(6.28) \quad \sum_{(a,A)=1} \frac{\mu^2(a)}{f_1(a)} \left(\sum_{s|p_1 \dots p_{r-1}} \mu(s) P^+ \left(\frac{\log R_2/as}{\log R_2} \right) \right)^2 = \frac{(\log R_2)^k}{\mathfrak{S}(\mathcal{L})(k-1)!} \int_0^1 G(1-t)^2 t^{k-1} dt + O((\log R_2)^{k-1}).$$

□

We also require a bound on the size of the sieve coefficients λ_d .

Lemma 6.7. *We have that*

$$\lambda_d \ll (\log R_2)^k.$$

Proof. This is proven in [3][Proof of Theorem 7].

□

We finish this section with a partial summation lemma, which will be useful later on.

Lemma 6.8. *Let $0 \leq a < b$ be fixed constants. Let $V : [a, b] \rightarrow \mathbb{R}_{\geq 0}$ be a continuous piecewise smooth function. If V satisfies $V(x) \ll x$ uniformly for $x \in [a, b]$ then we have*

$$\sum_{R^a \leq p \leq R^b} \frac{1}{p} V\left(\frac{\log p}{\log R}\right) = \int_a^b \frac{V(u)}{u} du + O\left(\frac{M(V) \log \log R}{\log R}\right),$$

where

$$M(V) = \sup_{t \in [a, b]} (1 + |V'(t)|).$$

Proof. The result follows straightforwardly by partial summation and the prime number theorem.

If $a = 0$ then we replace a with $2/\log R$. This leaves the left hand side of the result unchanged, and introduces an error

$$(6.29) \quad \int_a^{2/\log R} \frac{V(u)}{u} du \ll \frac{1}{\log R}$$

to the right hand side, which can be absorbed into the error term.

By the prime number theorem

$$(6.30) \quad \pi(y) = y \left(1 + O\left(\frac{1}{\log y}\right)\right).$$

Therefore, by partial summation we have

$$\begin{aligned} \sum_{R^a \leq p_j \leq R^b} \frac{1}{p} V\left(\frac{\log p}{\log R}\right) &= O\left(\frac{1}{\log R}\right) + \int_{R^a}^{R^b} \frac{t}{t^2 \log t} V\left(\frac{\log t}{\log R}\right) \left(1 + O\left(\frac{1}{\log t}\right)\right) dt \\ &\quad + \int_{R^a}^{R^b} \frac{t}{t^2 (\log t) (\log R)} V'\left(\frac{\log t}{\log R}\right) \left(1 + O\left(\frac{1}{\log t}\right)\right) dt \\ &= \int_a^b \frac{V(u)}{u} du + O\left(\int_a^b \frac{1 + |V'(u)|}{u \log R} du\right) + O\left(\frac{1}{\log R}\right) \\ (6.31) \quad &= \int_a^b \frac{V(u)}{u} du + O\left(\frac{M(V) \log \log R}{\log R}\right). \end{aligned}$$

□

7. PROOF OF PROPOSITION 5.1

We consider the weighted sum of Proposition 5.1 in a similar way to previous work on Selberg's $\Lambda^2 \Lambda^-$ sieve which in its basic form considers the weight $W_0(x) = -1$.

$$\begin{aligned}
 \sum_{N \leq n \leq 2N} \left(\sum_{\substack{p \mid \Pi(n) \\ p \leq R_1}} W_0 \left(\frac{\log p}{\log R_2} \right) \right) \left(\sum_{\substack{d \mid \Pi(n) \\ d \leq R_2}} \lambda_d \right)^2 &= \sum_{p \leq R_1} W_0 \left(\frac{\log p}{\log R_2} \right) \sum_{d, e \leq R_2} \lambda_d \lambda_e \sum_{\substack{N \leq n \leq 2N \\ [p, d, e] \mid \Pi(n)}} 1 \\
 &= N \sum'_{p \leq R_1} W_0 \left(\frac{\log p}{\log R_2} \right) \sum'_{d, e \leq R_2} \frac{\lambda_d \lambda_e}{f([d, e, p])} + O_{W_0}(E_1) \\
 (7.1) \quad &= N \sum'_{p \leq R_1} W_0 \left(\frac{\log p}{\log R_2} \right) \frac{T_p}{f(p)} + O_{W_0}(E_1),
 \end{aligned}$$

where

$$(7.2) \quad E_1 = \sum_{p \leq R_1} \sum_{d, e \leq R_2} |\lambda_d \lambda_e r_{[d, e, p]}|, \quad r_d = \sum_{\substack{N < n \leq 2N \\ d \mid \Pi(n)}} 1 - \frac{N}{f(d)}.$$

By Lemma 6.7 we have $\lambda_d \ll (\log N)^k$, and we note that $r_d \leq k^{\omega(d)}$. Therefore we have

$$\begin{aligned}
 E_1 &\ll (\log N)^{2k} \sum_{\substack{p \leq R_1 \\ d, e \leq R_2}} \mu^2([d, e, p]) k^{\omega([d, e, p])} \\
 &\ll (\log N)^{2k} \sum_{r \leq R_2^2 R_1} \mu^2(r) (7k)^{\omega(r)} \\
 &\ll (\log N)^{2k} R_2^2 R_1 \sum_{r \leq R_2^2 R_1} \frac{\mu^2(r) (7k)^{\omega(r)}}{r} \\
 &\ll (\log N)^{2k} R_2^2 R_1 \prod_{p \leq R_2^2 R_1} \left(1 + \frac{7k}{p} \right) \\
 (7.3) \quad &\ll (\log N)^{9k} R_2^2 R_1.
 \end{aligned}$$

Thus for $R_2^2 R_1 \leq N(\log N)^{-9k}$ we have $E_1 \ll N$.

By Lemma 6.6 we have

$$(7.4) \quad T_p = (\log R_2)^k \frac{\mathfrak{S}(\mathcal{L})}{(k-1)!} I_0 \left(\frac{\log p}{\log R_2} \right) + O((\log N)^{k-1} \log \log N),$$

where

$$(7.5) \quad I_0(x) = \int_0^1 (P_1^+(1-t) - P_1^+(1-t-x))^2 t^{k-1} dt.$$

Recalling that $f(p) = p/k$ for $p \nmid A$, we see that the error terms from T_p contribute

$$(7.6) \quad \ll_{W_0} (\log N)^{k-1} \log \log N \sum_{p \leq R_1} \frac{1}{p} \ll (\log N)^{k-1} (\log \log N)^2.$$

Therefore we are left to estimate the sum

$$(7.7) \quad \sum'_{p \leq R_1} \frac{1}{p} W_0 \left(\frac{\log p}{\log R_2} \right) I_0 \left(\frac{\log p}{\log R_2} \right).$$

We note that if $t \leq 1 - x$ then $P^+(1 - t) - P^+(1 - t - x) \ll x$, and so

$$(7.8) \quad I_0(x) \ll x.$$

If $1 - x \leq t \leq 1$ then since the interval has length x we also have

$$(7.9) \quad I_0(x) \ll x.$$

By the piecewise smoothness of $I_0(x)$ and $W_0(x)$ we have uniformly for $x \in [0, r1/r2]$

$$(7.10) \quad I'_0(x) \ll 1, \quad W'_0(x) \ll_{W_0} 1.$$

Therefore by Lemma 6.8, we have

$$(7.11) \quad \sum_{p \leq R_1} \frac{1}{p} W_0\left(\frac{\log p}{\log R_2}\right) I_0\left(\frac{\log p}{\log R_2}\right) = \int_0^{r_1/r_2} \frac{W_0(u)}{u} I_0(u) du + O_{W_0}\left(\frac{\log \log N}{\log N}\right).$$

By (7.8) we see that the contribution to the above sum for primes which divide A is

$$(7.12) \quad \ll \frac{1}{\log N}.$$

This gives the result.

8. PROOF OF PROPOSITION 5.2

We will follow a similar argument to that of Graham, Goldston, Pintz and Yıldırım[3] where the result was obtained with $r = 2$ and $W_2(x_1, x_2) = 1$. Thorne [11] extended this in the natural way to consider $r > 2$, again without the weighting W_r . In order to introduce the weighting by W_r , it is necessary to establish a Bombieri-Vinogradov style result for numbers with r prime factors weighted by W_r .

Lemma 8.1. *Let*

$$\beta_r(n) = \begin{cases} W_r\left(\frac{\log p_1}{\log n}, \dots, \frac{\log p_{r-1}}{\log n}\right), & n = p_1 p_2 \dots p_r \text{ with } p_1 \leq \dots \leq p_r, \\ 0, & \text{otherwise,} \end{cases}$$

for some piecewise smooth function $W_r : [0, 1]^{r-1} \rightarrow \mathbb{R}$.

Put

$$(8.1) \quad \Delta_{\beta,r}(x; q) = \max_{y \leq x} \max_{\substack{a \\ (a, q) = 1}} \left| \sum_{\substack{y < n \leq 2y \\ n \equiv a \pmod{q}}} \beta_r(n) - \frac{1}{\phi(q)} \sum_{\substack{y < n \leq 2y \\ (n, q) = 1}} \beta_r(n) \right|$$

For every fixed integer $h > 0$, and for every $C > 0$ there exists a constant $C' = C'(C, h)$ such that if $Q \leq x^{1/2}(\log x)^{-C'}$ then we have

$$(8.2) \quad \sum_{q \leq Q} \mu^2(q) h^{\omega(q)} \Delta_{\beta,r}(x; q) \ll_{C, h, W_r} x(\log x)^{-C}.$$

Proof. This result follows from the Bombieri-Vinogradov theorem for numbers with exactly r prime factors, as proven by Motohashi [8], and the continuity of W_r .

We assume that W_r is smooth. The result can be extended to piecewise smooth functions by taking smooth approximations.

We fix a constant $C > 0$, an integer h , and a function W_r .

We let

$$(8.3) \quad \chi_{\delta, \eta}(n) = \begin{cases} 1, & n = p_1 p_2 \dots p_r \text{ with } n^{\eta_i} \leq p_i \leq n^{\delta_i} \forall i \\ & \text{and } p_1 < p_2 < \dots < p_r. \\ 0, & \text{otherwise.} \end{cases}$$

By Motohashi's result [8][Theorem 2] we have that uniformly for any choice of constants δ_i and η_i ($i = 1, \dots, r$) there is a constant $C' = C'(C, h)$ such that if $Q \leq x^{1/2}(\log x)^{-C'}$ then we have

(8.4)

$$\sum_{q \leq Q} \mu^2(q) h^{\omega(q)} \max_{\substack{y, a \\ y \leq x \\ (a, q)=1}} \left| \sum_{\substack{y \leq n \leq 2y \\ n \equiv a \pmod{q}}} \chi_{\delta, \eta}(n) - \frac{1}{\phi(q)} \sum_{\substack{y \leq n \leq 2y \\ (n, q)=1}} \chi_{\delta, \eta}(n) \right| \ll_{C, h} x(\log x)^{-(C+h)(r+1)}.$$

We choose $\delta_i \in \{(\log x)^{-C-h}, 2(\log x)^{-C-h}, \dots, \lceil (\log x)^{C+h} \rceil (\log x)^{-C-h}\}$ separately for each $i \in \{1, \dots, r\}$, subject to the constraint $\delta_i \leq \delta_{i+1}$ ($1 \leq i \leq r-1$). For each choice of the δ_i we take $\eta_i = \delta_i - (\log x)^{-C-h}$ for $1 \leq i \leq r$. We put

$$(8.5) \quad W_r(\delta) = W_r(\delta_1, \delta_2, \dots, \delta_{r-1}).$$

We notice that by the smoothness of W_r we have that

$$(8.6) \quad \begin{aligned} \beta_r(n) &= \sum_{\delta} \chi_{\delta, \eta}(n) (W_r(\delta) + O((\log x)^{-C-h})) \\ &= \sum_{\delta} \chi_{\delta, \eta}(n) W_r(\delta) + O((\log x)^{-C-h}). \end{aligned}$$

Here \sum_{δ} indicates a sum over all the $O((\log x)^{r(C+h)})$ possible choices of the δ_i .

Therefore we have that

$$(8.7) \quad \begin{aligned} \sum_{\substack{y \leq n \leq 2y \\ n \equiv a \pmod{q}}} \beta_r(n) - \frac{1}{\phi(q)} \sum_{\substack{y \leq n \leq 2y \\ (n, q)=1}} \beta_r(n) &= \sum_{\delta} W_r(\delta) \left(\sum_{\substack{y \leq n \leq 2y \\ n \equiv a \pmod{q}}} \chi_{\delta, \eta}(n) - \frac{1}{\phi(q)} \sum_{\substack{y \leq n \leq 2y \\ (n, q)=1}} \chi_{\delta, \eta}(n) \right) \\ &\quad + O\left((\log y)^{-C-h} \frac{y}{\phi(q)}\right). \end{aligned}$$

Thus for $Q \leq x(\log x)^{-C'}$ we have

$$(8.8) \quad \begin{aligned} &\sum_{q \leq Q} \mu^2(r) h^{\omega(q)} \Delta_{\beta, r}(x; q) \\ &\leq \sum_{\delta} W_r(\delta) \sum_{q \leq Q} \mu^2(r) h^{\omega(q)} \max_{\substack{a, y \\ y \leq x \\ (a, q)=1}} \left| \sum_{\substack{y \leq n \leq 2y \\ n \equiv a \pmod{q}}} \chi_{\delta, \eta}(n) - \frac{1}{\phi(q)} \sum_{\substack{y \leq n \leq 2y \\ (n, q)=1}} \chi_{\delta, \eta}(n) \right| \\ &\quad + O\left((\log x)^{-(C+h)} \sum_{q \leq Q} \mu^2(q) h^{\omega(q)} \frac{x}{\phi(q)}\right) \\ &\ll \sum_{\delta} W_r(\delta) x(\log x)^{-(C+h)(r+1)} + x(\log x)^{-(C+h)} \prod_{p \leq Q} \left(1 + \frac{h}{p-1}\right) \\ &\ll x(\log x)^{-C}. \end{aligned}$$

□

With this, we can adapt the argument of Thorne [11] slightly to rewrite the main term in terms of the quantities T_q^* .

Lemma 8.2. *We have*

$$\sum_{N \leq n \leq 2N} \beta_r(L_j(n)) \left(\sum_{d|\Pi(n)} \lambda_d \right)^2 = \frac{AN}{\phi(A)(\log N)} \sum_{\substack{p_1, \dots, p_{r-1} \\ N^\epsilon < p_1 < p_2 < \dots < p_{r-1} \\ q < \min(N/R_2, N/p_{r-1})}} \frac{T_q^*}{q} \alpha \left(\frac{\log p_1}{\log R_2}, \dots, \frac{\log p_{r-1}}{\log R_2} \right) \\ + O_{W_r} \left(N(\log N)^{k-1} (\log \log N)^{r-1} \right),$$

where

$$q = \prod_{i=1}^{r-1} p_i, \\ T_\delta^* = \sum_{\substack{d, e \\ (d, A) = (e, A) = 1}} \frac{\lambda_d \lambda_e}{f^*([d, e, \delta]/\delta)}, \\ \alpha(q) = \left(\frac{\log N}{\log N - \log q} \right) W_r \left(\frac{\log p_1}{\log R_2}, \dots, \frac{\log p_{r-1}}{\log R_2} \right).$$

Proof. Thorne [11] considers essentially the same sum but without the weighting by W_r . In his argument up until equation (4.14) on Page 15, this difference only affects the argument when he appeals to the Bombieri-Vinogradov theorem for E_h numbers (where $h \leq r$). Lemma 8.1 gives the equivalent Bombieri-Vinogradov style result when weighting by W_r , and so exactly the same argument follows through. The only additional assumption of Thorne is that he restricts the consideration to numbers $n = p_1 \dots p_r$ satisfying

$$(8.9) \quad \exp(\sqrt{\log N}) < p_1 < \dots < p_r \quad \text{and} \quad R_2 < p_r.$$

This is satisfied if for a fixed $\epsilon > 0$ we require W_r to be supported on

$$(8.10) \quad \mathcal{A}_r = \left\{ x \in [0, 1]^{r-1} : \epsilon < x_1 < \dots < x_{r-1}, \sum_{i=1}^{r-1} x_i < \min(1 - r_2, 1 - x_{r-1}) \right\}.$$

This gives us in our case (the equivalent of Thorne's equation (4.14) but with the explicit error term he calculates)

$$\sum_{N \leq n \leq 2N} \beta_r(L_j(n)) \left(\sum_{d|\Pi(n)} \lambda_d \right)^2 = \sum'_{d, e} \lambda_d \lambda_e \sum_{p_1, \dots, p_{r-1}}^* \frac{d_{k-1}([d, e, q]/q)}{\phi(a_j[d, e, q]/q)} \\ \times \sum_{a_j N/q \leq m \leq 2a_j N/q} 1_{\mathbb{P}}(n) W_r \left(\frac{\log p_1}{\log mq}, \dots, \frac{\log p_{r-1}}{\log mq} \right) \\ + O(N). \quad (8.11)$$

Here and from now we use the symbol \sum^* to indicate that we are summing over primes p_1, \dots, p_{r-1} with

$$(8.12) \quad \left(\frac{\log p_1}{\log N}, \dots, \frac{\log p_{r-1}}{\log N} \right) \in \mathcal{A}_r.$$

Again we assume for simplicity that W_r is smooth. By taking smooth approximations one can establish the result for piecewise-smooth W_r .

Estimating the inner sum gives

$$\begin{aligned}
 & \sum_{a_j N/q \leq m \leq 2a_j N/q} 1_{\mathbb{P}}(m) W_r \left(\frac{\log p_1}{\log m q}, \dots, \frac{\log p_{r-1}}{\log m q} \right) \\
 &= \left(W_r \left(\frac{\log p_1}{\log N}, \dots, \frac{\log p_{r-1}}{\log N} \right) + O \left(\frac{1}{\log N} \right) \right) \left(\pi \left(\frac{2a_j N}{q} \right) - \pi \left(\frac{a_j N}{q} \right) \right) \\
 (8.13) \quad &= W_r \left(\frac{\log p_1}{\log N}, \dots, \frac{\log p_{r-1}}{\log N} \right) \frac{a_j N}{\log N} \left(\frac{\log N}{\log N - \log q} \right) \left(1 + O \left(\frac{1}{\log N} \right) \right).
 \end{aligned}$$

We note that by Hypothesis 1 if $d \mid \Pi(n)$ then $(d, A) = 1$. Therefore $(a_j, [d, e, q]/q) = 1$, so $\phi(a_j [d, e, q]/q) = \phi(a_j) \phi([d, e, q]/q)$. Together these give

$$\begin{aligned}
 & \sum_{N \leq n \leq 2N} \beta_r(L_j(n)) \left(\sum_{d \mid \Pi(n)} \lambda_d \right)^2 \\
 &= \frac{a_j N}{\phi(a_j) \log N} \sum_{p_1, \dots, p_{r-1}}^* \frac{T_q^* W_r \left(\frac{\log p_1}{\log N}, \dots, \frac{\log p_{r-1}}{\log N} \right) \log N}{q(\log N - \log q)} (1 + O((\log N)^{-1})) \\
 &\quad + O(N) \\
 &= \frac{a_j N}{\phi(a_j) \log N} \sum_{p_1, \dots, p_{r-1}}^* \frac{T_q^*}{q} \alpha \left(\frac{\log p_1}{\log R_2}, \dots, \frac{\log p_{r-1}}{\log R_2} \right) \left(1 + O \left(\frac{1}{\log N} \right) \right) \\
 (8.14) \quad &\quad + O(N),
 \end{aligned}$$

where

$$(8.15) \quad \alpha(x_1, \dots, x_{r-1}) = \frac{W_r(r_2 x_1, \dots, r_2 x_{r-1})}{1 - r_2 \sum_{i=1}^{r-1} x_i}.$$

We note that a_j and A are composed of the same prime factors, so $a_j/\phi(a_j) = A/\phi(A)$. Therefore the main term is that of the Lemma.

By Lemma 6.5 we have

$$(8.16) \quad T_q^* \ll_r (\log N)^k \log p_1 + (\log N)^k \log \log N.$$

We also have

$$(8.17) \quad \alpha(x_1, \dots, x_{r-1}) \ll_{W_r} 1.$$

Thus the $O(1/\log N)$ term contributes

$$(8.18) \quad \ll_{W_r, r} N(\log N)^{k-2} \sum_{p_1, \dots, p_{r-1}}^* \frac{\log p_1 + \log \log N}{p_1 \dots p_{r-1}} \ll_{W_r, r} N(\log N)^{k-1} (\log \log N)^{r-2}.$$

This gives the result. \square

Lemma 8.3. *We have*

$$\begin{aligned}
 & \sum_{p_1, \dots, p_{r-1}}^* \frac{T_q^*}{q} \alpha \left(\frac{\log p_1}{\log R_2}, \dots, \frac{\log p_{r-1}}{\log R_2} \right) \\
 &= (\log R_2)^{k+1} \frac{\phi(A) \mathfrak{S}(\mathcal{L})}{A(k-2)!} \int \dots \int \frac{I_1(u_1, \dots, u_{r-1}) \alpha(u_1, \dots, u_{r-1})}{u_1 u_2 \dots u_{r-1}} du_1 \dots du_{r-1} \\
 (8.19) \quad &+ O((\log \log N)^r (\log N)^k)
 \end{aligned}$$

Where the integration is subject to the constraints

$$(8.20) \quad \epsilon < u_1 < \dots < u_{r-1}, \quad \text{and} \quad \sum_{i=1}^{r-1} u_i \leq \min(r_2^{-1} - 1, r_2^{-1} - u_{r-1}).$$

Proof. By Lemma 6.6 for $q = p_1 p_2 \dots p_{r-1}$ we have

$$(8.21) \quad T_q^* = (\log R_2)^{k+1} \frac{\phi(A) \mathfrak{S}(\mathcal{L})}{A(k-2)!} I_1 \left(\frac{\log p_1}{\log R_2}, \dots, \frac{\log p_{r-1}}{\log R_2} \right) + O_r((\log N)^k \log \log N).$$

Thus summing the error term over p_1, \dots, p_{r-1} gives a contribution

$$\begin{aligned}
 & \sum_{p_1, \dots, p_{r-1}}^* \frac{1}{q} \alpha \left(\frac{\log p_1}{\log R_2}, \dots, \frac{\log p_{r-1}}{\log R_2} \right) (\log N)^k \log \log N \\
 & \ll_{W_r} (\log N)^k \log \log N \left(\sum_{p \leq N} \frac{1}{p} \right)^{r-1} \\
 (8.22) \quad & \ll_{W_r} (\log N)^k (\log \log N)^r.
 \end{aligned}$$

We are therefore left to evaluate the main term

$$(8.23) \quad \sum_{p_1, \dots, p_{r-1}}^* \frac{1}{q} \alpha \left(\frac{\log p_1}{\log R_2}, \dots, \frac{\log p_{r-1}}{\log R_2} \right) I_1 \left(\frac{\log p_1}{\log R_2}, \dots, \frac{\log p_{r-1}}{\log R_2} \right).$$

We will now apply Lemma 6.8 to p_{r-1}, \dots, p_1 in turn to estimate the sum $\sum^* \alpha(q) T_q^* q^{-1}$.

For $u_1, \dots, u_j \in [0, r_2^{-1}]$ we put

$$(8.24) \quad V_j(u_1, \dots, u_j) = \int \dots \int \frac{1}{\prod_{i=j+1}^{r-1} u_i} \alpha(u_1, \dots, u_{r-1}) I_1(u_1, \dots, u_{r-1}) du_{j+1} \dots du_{r-1},$$

where the integration is subject to $u_j < u_{j+1} < \dots < u_{r-1}$ and $\sum_{i=1}^{r-1} u_i \leq \min(r_2^{-1} - 1, r_2^{-1} - u_{r-1})$.

As in the proof of Lemma 6.5, since \tilde{P} is continuous and its derivative is uniformly bounded on $[0, 1]$, we have that

$$\begin{aligned}
 I_1(u_1, \dots, u_{r-1}) &= \int_0^1 \left(\sum_{J \subset \{1, \dots, r-1\}} \tilde{P}^+ \left(1 - t - \sum_{i \in J} u_i \right) (-1)^{|J|} \right)^2 t^{k-2} dt \\
 &\ll_r \int_0^1 \sum_{J \subset \{1, \dots, r-1\} \setminus \{j\}} \left(\tilde{P}^+ \left(1 - t - \sum_{i \in J} u_i \right) - \tilde{P}^+ \left(1 - t - u_j - \sum_{i \in J} u_i \right) \right)^2 t^{k-2} dt \\
 (8.25) \quad &\ll_r u_j^2.
 \end{aligned}$$

Thus, since $\alpha(u_1, \dots, u_{r-1}) \ll 1$, we have uniformly for $u_1, \dots, u_j \in [0, r_2^{-1}]$

$$\begin{aligned} V_j(u_1, \dots, u_j) &\ll u_j^2 \int \dots \int \frac{1}{\prod_{i=j+1}^{r-1} u_i} du_{j+1} \dots du_{r-1} \\ &\ll u_j^2 (1 + |\log 1/u_j|^r) \\ (8.26) \quad &\ll u_j. \end{aligned}$$

Moreover, essentially the same argument shows that uniformly for $u_1, \dots, u_j \in [0, r_2^{-1}]$ we have

$$(8.27) \quad \frac{\partial}{\partial u_j} I_1(u_1, \dots, u_{r-1}) \ll_r u_j.$$

Thus since

$$(8.28) \quad \frac{\partial}{\partial u_j} \alpha(u_1, \dots, u_{r-1}) \ll_r 1$$

we have that

$$\begin{aligned} \frac{\partial}{\partial u_j} V_j(u_1, \dots, u_j) &\ll u_j \int \dots \int \frac{1}{\prod_{i=j+1}^{r-1} u_i} du_{j+1} \dots du_{r-1} \\ (8.29) \quad &\ll 1. \end{aligned}$$

Thus the condition of Lemma 6.8 applies for the function V_j . Applying Lemma 6.8 in turn to $V_{r-1}, V_{r-2}, \dots, V_1$ gives the result. We note that the error terms contribute a total which is $\ll (\log N)^k (\log \log N)^{r-1}$. \square

9. PROOF OF PROPOSITION 5.3

By Lemma 6.7 we have $\lambda_d \ll (\log N)^k$. Therefore we have

$$\begin{aligned} \sum_{p \leq AN^{1/2}} \sum_{\substack{N \leq n \leq 2N \\ p^2 | \Pi(n)}} \left(\sum_{\substack{d | \Pi(n) \\ d \leq R_2}} \lambda_d \right)^2 &= N \sum_{p \leq AN^{1/2}} \sum'_{d, e \leq R_2} \frac{\lambda_d \lambda_e}{f([d, e, p^2])} + O \left(\sum_{p \leq AN^{1/2}} \sum_{d, e \leq R_2} |\lambda_d \lambda_e r_{[d, e, p^2]}| \right) \\ (9.1) \quad &\ll N \frac{T_p}{p^2} + O \left((\log N)^{2k} \sum_{r \leq R_2^2 AN^{1/2}} \mu^2(r) (7k)^{\omega(r)} \right). \end{aligned}$$

We first bound the error term

$$\begin{aligned} \sum_{r \leq R_2^2 AN^{1/2}} \mu^2(r) (7k)^{\omega(r)} &\ll R_2^2 N^{1/2} \sum_{r \leq AR_2^2 N^{1/2}} \frac{\mu^2(r) (7k)^{\omega(r)}}{r} \\ &\ll R_2^2 N^{1/2} \prod_{p \leq AR_2^2 N^{1/2}} \left(1 + \frac{7k}{p} \right) \\ (9.2) \quad &\ll R_2^2 N^{1/2} (\log N)^{7k}. \end{aligned}$$

Thus for $R_2 \leq N^{1/4} (\log N)^{-5k}$ the error term is $O(N)$.

By Lemma 6.5 we have that

$$(9.3) \quad T_p \ll (\log N)^{k-1} \log p + (\log N)^{k-1} \log \log N.$$

Thus

$$\begin{aligned}
 \sum_{p \leq AN^{1/2}} \sum_{\substack{N \leq n \leq 2N \\ p^2 \mid \Pi(n)}} \left(\sum_{\substack{d \mid \Pi(n) \\ d \leq R_2}} \lambda_d \right)^2 &\ll N(\log N)^{k-1} \sum_{p \leq N^{1/4}} \frac{\log p + \log \log N}{p^2} + O(N) \\
 (9.4) \qquad \qquad \qquad &\ll N(\log N)^{k-1} \log \log N.
 \end{aligned}$$

10. ACKNOWLEDGMENT

I would like to thank my supervisor, Prof. Heath-Brown and Dr. Craig Franze for many helpful comments.

REFERENCES

- [1] CHEN, J. R. On the representation of a larger even integer as the sum of a prime and the product of at most two primes. *Sci. Sinica* 16 (1973), 157–176.
- [2] DIAMOND, H., AND HALBERSTAM, H. Some applications of sieves of dimension exceeding 1. In *Sieve methods, exponential sums, and their applications in number theory (Cardiff, 1995)*, vol. 237 of *London Math. Soc. Lecture Note Ser.* Cambridge Univ. Press, Cambridge, 1997, pp. 101–107.
- [3] GOLDSTON, D. A., GRAHAM, S. W., PINTZ, J., AND YILDIRIM, C. Y. Small gaps between products of two primes. *Proc. Lond. Math. Soc.* (3) 98, 3 (2009), 741–774.
- [4] GOLDSTON, D. A., PINTZ, J., AND YILDIRIM, C. Y. Primes in tuples. II. *Acta Math.* 204, 1 (2010), 1–47.
- [5] HEATH-BROWN, D. R. Almost-prime k -tuples. *Mathematika* 44, 2 (1997), 245–266.
- [6] HO, K.-H., AND TSANG, K.-M. On almost prime k -tuples. *J. Number Theory* 120, 1 (2006), 33–46.
- [7] MAYNARD, J. 3-tuples have at most 7 prime factors infinitely often. *Pre-print*.
- [8] MOTOHASHI, Y. An induction principle for the generalization of Bombieri’s prime number theorem. *Proc. Japan Acad.* 52, 6 (1976), 273–275.
- [9] PORTER, J. W. Some numerical results in the Selberg sieve method. *Acta Arith.* 20 (1972), 417–421.
- [10] SELBERG, A. *Collected papers. Vol. II.* Springer-Verlag, Berlin, 1991. With a foreword by K. Chandrasekharan.
- [11] THORNE, F. Bounded gaps between products of primes with applications to ideal class groups and elliptic curves. *Int. Math. Res. Not. IMRN*, 5 (2008), Art. ID rnm 156, 41.

MATHEMATICAL INSTITUTE, 24-29 ST GILES’, OXFORD, OX1 3LB

E-mail address: maynard@math.ox.ac.uk